ITEA Journal 2009; 30: 421–433

Examining the Role of Metadata in Testing IED Detection Systems

Paul J. Fortier, Ph.D. and Kiran Dasari
University of Massachusetts Dartmouth, North Dartmouth, Massachusetts

A sensor provides capabilities to extract unique spatial and temporal instances of a measurement specific to a system under test. The measurement is only useful when the interrogation agent fully knows the measurement's context. After a measurement event, if insufficient descriptive, contextual information is available, the measurement is lost for use in ascertaining future value. This article defines generic concepts for metadata and sensors such that their design, selection, application, and use are fully described for users, allowing for seamless present or future uses. Within this context, we examine a definition of metadata, metadata types, metadata uses, concepts for sensor ontology, and multilevel sensor metadata. Metadata's role in detecting events; acquiring measurements; converting measurements to information, information fusion, and aggregation into complex structures; developing actionable knowledge and persistent storage; and retrieving derived knowledge are addressed through presentation of an illustrative example application, that of improvised explosive device detection system testing.

Key words: Metadata; sensors; domain context; measurement; information; actionable knowledge; knowledge storage and retrieval.

Localization, selection, extraction, processing, and movement of measurement data from a system under test (SUT) to a spatial and temporal data archive, as well as query selection and reuse for new analysis purposes, represents a challenging set of tasks. Significant effort is spent in translating raw measurements to useful information through the application of time and spatial associations, calibration details, and device pedigree rankings to name a few. Additionally, time and effort is spent adding further meaning to extracted information fragments to include data usage tracking, determining environmental and situational conditions relevant to the extracted measurement, as well as relevant contextual configurations (such as sensor settings, sensor base definitions, sensor placement, sensor pedigree, sensor orientation, etc.) to support formal assessment and interpretation of the raw measurement data. Information often is collected without regard for sensor health and is assumed correct and accurate. Such ad hoc methods are unacceptable if data are used in decision making or within human safety applications.

Data measurements represent raw atomic analog and or digital values derived from a discrete sampling device. Data alone provide little utility; consider a simple data set: *10 08 09 12 10*. The values could be interpreted in many ways; they could represent five distinct raw measurements for a single sensor (e.g., voltage levels), or they could represent a complex collection of grouped and organized items representing something totally different, such as a date (10-09-2008), time (1200 hours), and a temperature (10°C). To determine meaning requires that we add a relatively small amount of descriptive information to provide an accurate interpretation of the raw data. Without this descriptive information (metadata), there is little if any added utility to sampled and collected measurements.

The focus of this article is on the development of end-to-end metadata concepts aimed at derivation of actionable situational awareness for a target domain using information added to, extracted from, or derived from a collection of sensors. The sensors physically can be mobile or fixed, remote or local, appliqué or embedded, onboard or off board, dumb or smart. The primary driving requirement is the need to measure, collect, preserve, communicate, and share sensor-derived information for a variety of present and as yet unforeseen future test and evaluation applications. To accomplish this goal, we must make metadata

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
SEP 2009		**00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Examining the Role of Metadata in Testing IED Detection Systems	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Massachusetts Dartmouth,285 Old Westport Road ,North Dartmouth,MA,02747-2300**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**13**	

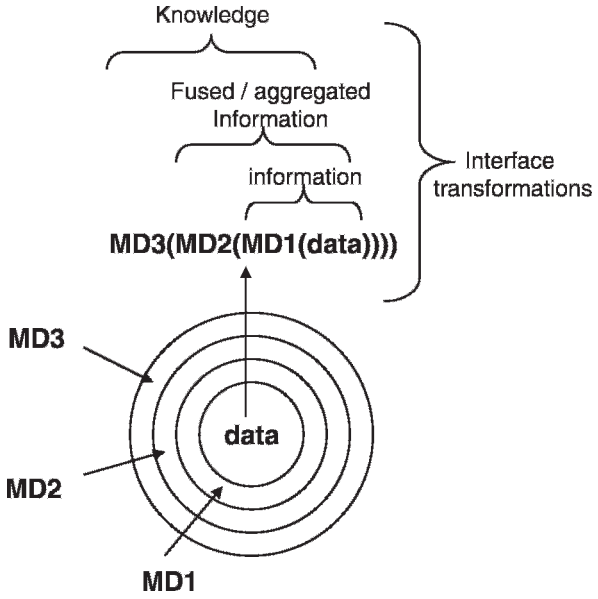

Figure 1. Metadata's role in data transformation.

available to aid in placing domain context and state to collected information to support information reuse or to support new initially unforeseen applications.

Metadata are the means through which raw data are transformed into information and are aggregated and fused into new synthetic representations, and finally through ontology into knowledge *(Figure 1)*. Metadata should aid in defining what data are present, what aspect of reality do the data represent, where the data are located, how the data got located there, how we acquire the data, how the data may be used, who may use the data, what available transformations or services are available to act upon this data, to name a few. The metadata management components for a test and evaluation (T&E) system should be able to clearly and concisely provide answers and direction to address each of these issues for all data and informational items under its management and control.

Metadata can be highly organized and formal, such as would be demonstrated by domain ontology (Borst 1997) or more ad hoc as found in an application program's data type specification. An ontology or context specification (Gauvin, Boury-Brisset, and Auger 2004) is the result of a study to categorize and organize items (Gruber 1993), that exist within some domain (Macintyre 1972). Such domain terminologies, optimized for human processing, are characterized by a significant amount of implicit knowledge. Deeper meaning for T&E can be achieved by constructing a T&E domain ontology. The benefits from such an exercise include: the ability to build more powerful and more interoperable information systems; support for the requirement to transmit, reuse, and share system and test data in real time and for future uses; semantic-based criteria to support different statistical aggregations; and possibly the most significant benefit an ontology brings to T&E systems is the ability to support the integration of knowledge and information (CCDA 2000; NISO 2004).

Metadata

Metadata has many uses, not simply to define information structure or domain transformations (Baca 1998; Duval et al. 2002). Metadata can describe how to process a collection of diverse information to represent some abstract construct (Bose 2002; Tambouris, Manouselis, and Costopoulou 2007) such as a synthetic or virtual sensor specification (Sethunadh, Athuladevi, and Iyer 2002). Metadata can describe system interaction, or how to convert information into knowledge (Ladner and Pe 2005). A requirement for the embedded instrumentation systems architecture (Michel and Fortier 2006, Visnevski 2008) was to generically define informational, control, and behavioral models to support embedded nonintrusive sensors in a T&E environment using metadata. Metadata can describe informational data flow, data aggregations, and data fusion supporting real-time synthetic sensor construction and operations (Sethunadh, Athuladevi, and Iyer, 2002; Visnevski and Johnson 2007).

Metadata classification

There is not one format or definition for metadata within the context of T&E systems applied to native, embedded, appliqué, or noncontact nonintrusive instrumentation (NII). Typical research and trade literature defines metadata according to structure and semantics (Qin and Prado 2006), or functions (services) supported (Tan 2004; Tannenbaum 1998). Within these two broad categories lie additional refinements for metadata classification. Upon review of the present research and trade literature, nine distinct classifications or types of metadata become apparent; these are as follows: descriptive, structural, administrative, preservation, usage, interface, transport, context, and process metadata. Each distinct type has a place within the T&E community and NII, and is briefly described.

Descriptive metadata are used to define, identify, and describe a measurement resource. The transducer electronic data sheets (TEDs) (Lee 2006b) and SensorML (Botts et al. 2004) represent a standardized digital means for providing specifications for sensor components and systems. SensorML's metadata includes identifiers, classifiers, constraints (time, legal, and security), capabilities, characteristics, contacts, and references in addition to inputs, outputs, parameters,

and system location, which can be mined and used for discovery of sensor systems and observation processes. TEDs provide metadata to aid in the definition of the sensor measurement device often referred to as the transducer, as well as definitions for basic elements and functional elements of a sensor using IEEE 1451's concept of a functional block.

Structural metadata provides information to define or organize complex collections of information items or to define a composite measurement composed of these basic information items (Hall and Llinas 1997). The definitions may include algorithms (possibly even a chain of algorithms needed to produce a desired derivative measurement) to use in the selection, extraction, fusion, aggregation, and combination of signal streams into a synthetic measurement or intermediate computation for use in a further informational refinement.

Administrative metadata represents information needed to manage and supervise all system metadata. Administrative metadata may include access rights information to metadata and data items, version control information for an item, data creator information, location information, and data production information supporting maintenance and supervision of data resources (Park et al. 2006).

Preservation metadata provides information needed to store and maintain information in a persistent, recoverable form within an archive. Preservation metadata may include physical assessment of data, the media they are housed in, storage formats, refresh rates, history of refresh, rebuild or recovery history, redundant copy location, repository or media status, pedigree, provenance, and other information related to long-term storage management (Ledlie et al. 2005).

Usage metadata maintains a record of how a data item and related metadata were utilized. One could look at this form of metadata as state maintenance or log history of operations concerning a data item. The metadata may include data inputs, outputs, intermediate values, and process chain descriptions along with configuration information to allow for restoration or reconstruction of a usage thread (Groth, Luc, and Moreau 2004).

Interface metadata define inputs to a stage or phase of operations for a system, along with functionality or possibly protocols applied to the input to transform them to the appropriate output format. Typical definitions may include which additional metadata are added to inputs so outputs can be transformed and formatted correctly to pass useful information to the next requesting level or phase of systems operations.

Transport metadata are an essential type of metadata when a communications media of any kind is associated with a domain. Transport metadata are used to define the payload (packet, stream, etc.) format for a transmission protocol and the transmission protocol (e.g., Transmission Control Protocol (TCP), User Datagram Protocol (UDP), etc.) steps or traces that should be maintained to reconstruct access patterns. Transport metadata may also maintain information concerning quality of service and other network parameters (e.g., addressing formats, integrity of payload, correctness, etc.) to allow transported data to be extracted correctly and made available for use (Faulstich and Grace 2007). The Transducer ML standard defines a self-describing data exchange protocol and common metadata data format standard based on XML supporting data streaming between any sensor and a processing sink (TML 2008) (Havens 2007).

Context metadata refers to complex relationship-oriented information (Borst 1997; Bose 2002), concerning how data relate to each other and under what conditions these relationships hold, such as found in an ontological representation of a domain's knowledge. Context may include fundamental classification of items, the provenance of complex items, pedigree relationships for items, and possibly even definitions for use cases for instances of measurement and event classes in the domain.

Process metadata describe information concerning the behavior and interface of processes and workloads (Nle et al. 2006). Metadata describing processes, algorithms, and methods could be stored in an object code or interface library for reuse by different process model instances in another domain. The TENA object model (Noseworthy 2005) and LINUX binaries represent such forms of process metadata.

Uses of metadata

Just as there are multiple types of metadata, there are also many differing uses. Research literature reveals nine general categories defining metadata uses: archival and preservation, digital identification, resource discovery, e-resource organization and management, observational retrieval, operational logging, interoperability management, knowledge discovery, and application development.

Archival and preservation metadata provides instructional information to aid in storing, recovering, restoring, locating, describing media, lifecycle management, error handling, and assessing physical status of SUT data measurements.

Digital identification services aid in appropriately identifying a physical device. Metadata services allow for new device type determination, localization, categorization (e.g., provenance and pedigree), and description of basic features of a device in relation to existing devices.

Resource discovery metadata services provide for the request and discovery of embedded, appliqué, and test devices, including those over a network if applicable. Services to maintain directories of active and known sensors as well as interface services to provide for plug-and-play access of sensors support resource discovery.

E-resource organization and management metadata services keep track of sensors once discovered or inserted into a SUT. This could include specification and build of a synthetic sensor from existing sensor and processing inventories or through selection and linking of sensors and possibly external processing to form a new as-yet-unimagined synthetic sensor. Of interest to E-resource metadata are management oriented tasks such as sensor configuration, reconfiguration, availability assessment, calibration, and coordination, such as sensor subscription and automeasurement retrieval.

Observation retrieval is a basic service of sensors and requires metadata to support registered sensors and users (sources and sinks), automatic pushing of observed measurements, as well as ad hoc pulling of measurements based on user requests. Translations may require specialized translation algorithms and accuracy parameter (or lineage) metadata.

Operational logging is a primary user of metadata supporting pedigree and provenance associations for measurements and requires capabilities to trace all forms of actions within a system.

Interoperability has many meanings, and in general provides procedures, processes, policies, and mechanisms to support the use of an item developed and applied in one domain to another possibly unanticipated domain. Dependency on the degree of seamless operations required will dictate the amount of metadata needed to provide for data or functional interoperability. The basic idea is to provide metamodels allowing for code to run independent of a platform (hardware, operating system, language, etc.) or data requirements.

Knowledge discovery utilizes all forms of domain and context information provided through domain ontology and a resultant use case bases describing instances of the domain knowledge. Metadata linking these case bases and ontological metadata stores are needed to enhance knowledge discovery.

Applications development is typically not thought of as a fundamental user of metadata. However, in sensor networks for T&E there will be a need for a variety of services for application development to maintain, use, or generate metadata concerning programming interface (application programming interfaces, graphical user interfaces) system configuration, sensor status (including fault assessment, management, reconfiguration, correction, etc.), and workflow configurations.

Example systems use of metadata

Sensors and sensor networks were originally driven by military and security concerns, and are now being developed and fielded into previously unenvisioned applications (e.g., habitat and environmental monitoring, pollution assessment, renewable energy management, home energy management, crop assessment, weather forecasting, disaster alerts, and endangered species assessment [Biagioni and Bridges 2002; Mainwaring et al. 2004;Wang 2003]).

Recent efforts in sensor networks focus on the need to refine and standardize on interoperable services (Lee 2006a; Lee and Percivall 2008). Services can roughly be broken into three categories: sensor management, operational development services, and applications development services. Sensor management services define policies and mechanisms for using sensor operational services to localize, configure, control access, and manage operations of sensor networks. Sensor application development services focus on services and development tools aiding sensor definition, integration, configuration management, fault management, and system assessment. Services for sensor operations support sensor specification, interface specification, directory management, collection services, observational services, notification services, planning services, coordination services, transactional services, information aggregation and fusion services, persistent storage and archival services, operational logging services, and configuration services (Lee 2006b).

One integration effort generalizing sensor networks architecture and services is Embedded Instrumentation System Architecture (EISA) (Michel and Fortier 2006; Visnevski 2008). EISA is an initiative funded under the Test Resource Management Center (TRMC) T&E/Science and Technology (S&T) NII focus area and has as a fundamental objective to develop a common, comprehensive methodology for nonintrusively collecting massive amounts of T&E data supporting war fighter systems testing. EISA offers a metadata driven methodology and common architecture for heterogeneous data collection, aggregation, and fusion in a real-time synchronized and correlated fashion. These methods support the real-time instrumentation and sensor management supporting nonintrusive synthetic (virtual) and real test instrumentation.

Initial sensor services mappings to the EISA architectural framework *(Figure 2)* indicates metadata related to defining resources, managing sensor operational conditions, calibrating sources, and configuring collection and transport services are found in the lowest three levels of the Information Technical Reference Model (ITRM) pyramid (Joshi and Michel 2007).

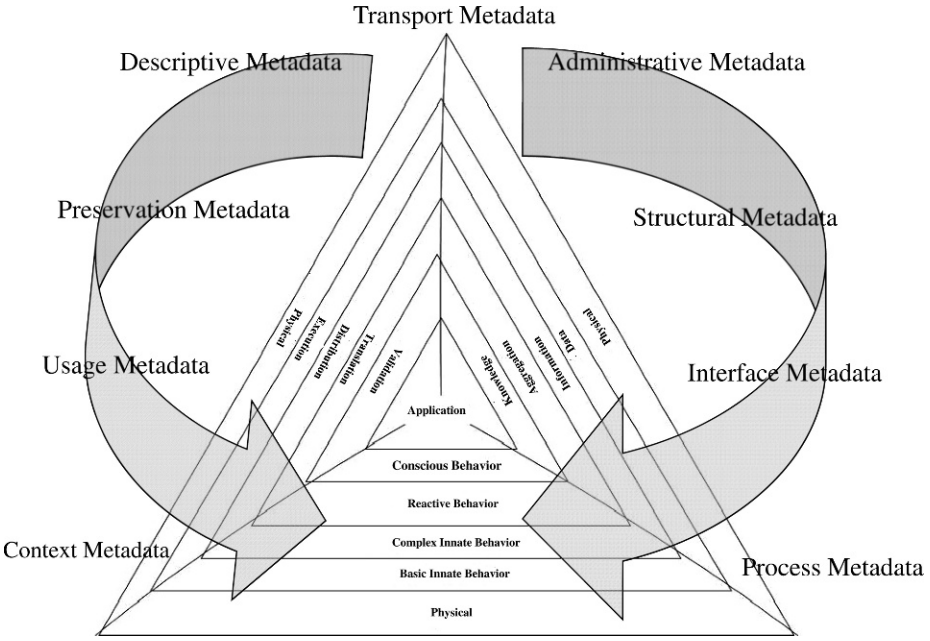

Figure 2. EISA and metadata mapping.

Metadata related to information aggregation and fusion, context refinement, scenario mapping, knowledge representation, and user administration are found in the upper layers. Metadata services such as test planning, measurement observation acquisition, alert configuration, and management utilizing myriad forms of metadata likewise map to numerous layers. EISA and the ITRM provide evolving templates to develop a mapping of metadata types and services for sensors applied to disparate domains such as medical informatics and environmental sustainability monitoring (Dasari 2008) as well as military and homeland security systems.

IED detection system T&E

For purposes of this article, the example focuses on metadata generated, extracted, or derived to construct and perform improvised explosive device (IED) detection systems testing within a generic test range *(Figure 3)*. Testing of IED detection products will be ongoing and evolving as the enemy's tactics and technology evolve. Each service has a number of products, projects, and proposed products in some stage of research and development. Such programs include Buckeye, developed jointly with the Army Corps of Engineers (Kauchak 2006) using imagery analysis; Shadow (Harpel 2007), an Army unmanned

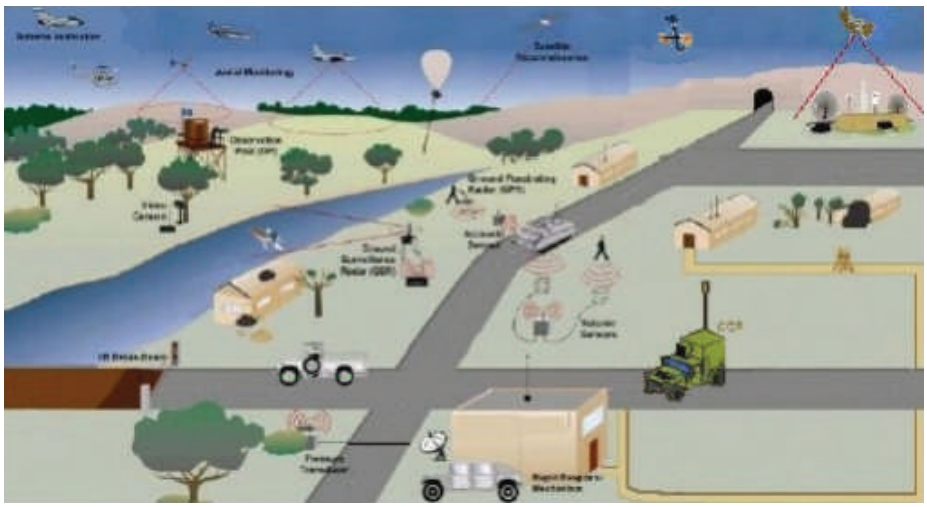

Figure 3. Test range for IED detection and defeat systems evaluation.

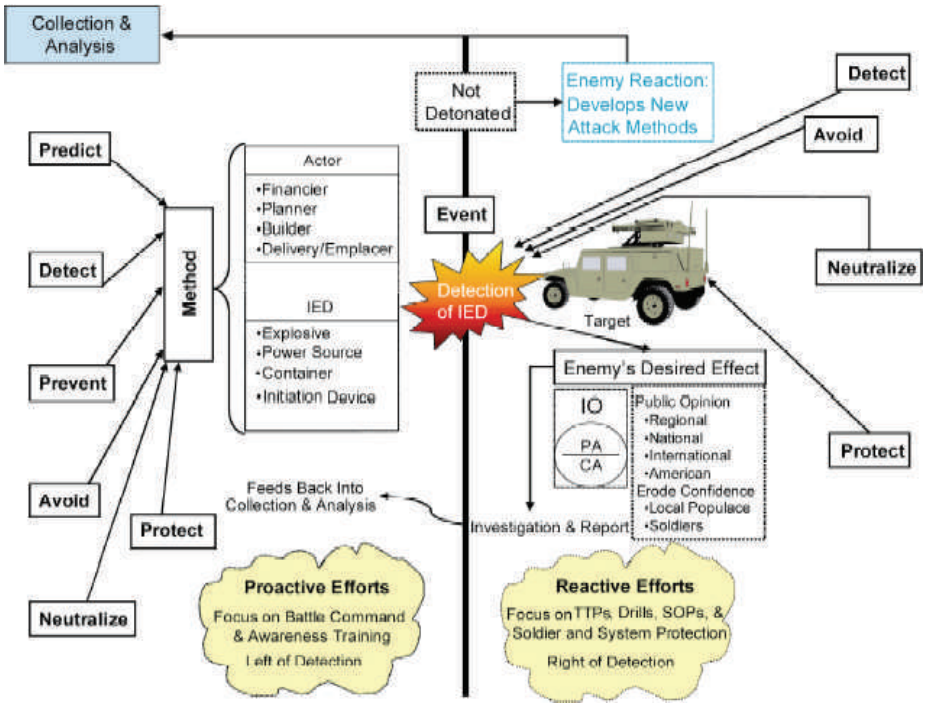

Figure 4. IED defeat system components.

autonomous vehicle (UAV) system using change detection and hyper spectral sensing; British Aerospace (BAE) systems' Talon Radiance II hyper spectral sensor system using both change detection and data mining technology to detect possible IED sites; EDO's Joint Counter Radio-Controlled Electronic Warfare (JCREW IED) jammer system; Northrop Grumman's Vehicle and Dismount Exploitation Radar (VADER) system using radar mounted on an airborne UAV system; and the Israeli Trophy Active Defense System (ADS) examined by the Army for fielding in Iraq.

The IED as a weapon has been known for decades in various forms (land mines, booby traps, suicide bomber, Kamikaze, etc.). It is only recently because of the conflicts in Iraq and Afghanistan that the Military has begun a rigorous examination of IEDs as a serious and coordinated threat and weapon. The Department of Defense (DOD) has begun development of technology for detecting a wide variety of these devices as part of a coordinated Joint IED detection and defeat organization. Typically IED devices are fielded in one of three primary forms: vehicle borne IED, suicide bomber (person-borne IED), and ad hoc munitions or Leave behind IED (e.g,. a C4 charge implanted in an animal carcass with cell phone detonator thrown on the side of a road, or as a pipe bomb taped to a target).

A complete anti-IED system requires an IED detection component, an IED assessment component, and an IED defeat component *(Figure 4)*. The IED detection system test example developed is limited to the testing of anti-IED detection hardware, software, procedures, metadata, and information during the arming, detonation, and assessment periods within the lifecycle of an IED *(Figure 5)*. Not included is the testing of anti-IED support platforms (such as a mine resistant ambush protection (MRAP) or joint EOD rapid response vehicle (JERRV), human operators, and countermeasures, though testing that includes such platforms as part of an environment for scoring and assessment of sensors and or procedures performance are considered. IEDs are constructed from a variety of elements *(Figure 6)*, all of which must be represented in a testing environment (typically not in their active form) for use in testing IED detection components.

Example scenario

In the IED detection system test scenario, we examine three phases of system T&E: the test planning stage, the test preparation phase, and the test execution phase. In each of these test phases, we examine the metadata requirements, existing tools, and techniques that exist to provide the needed metadata and services as well as define shortfalls and issues lacking.

In the scenario, a test engineer wishes to test the war fighting worthiness of a new IED distributed multi-sensor detection system. The test engineer needs to design a global situational assessment Data Acquisition System (DAS) using legacy, manufacturer, and newly

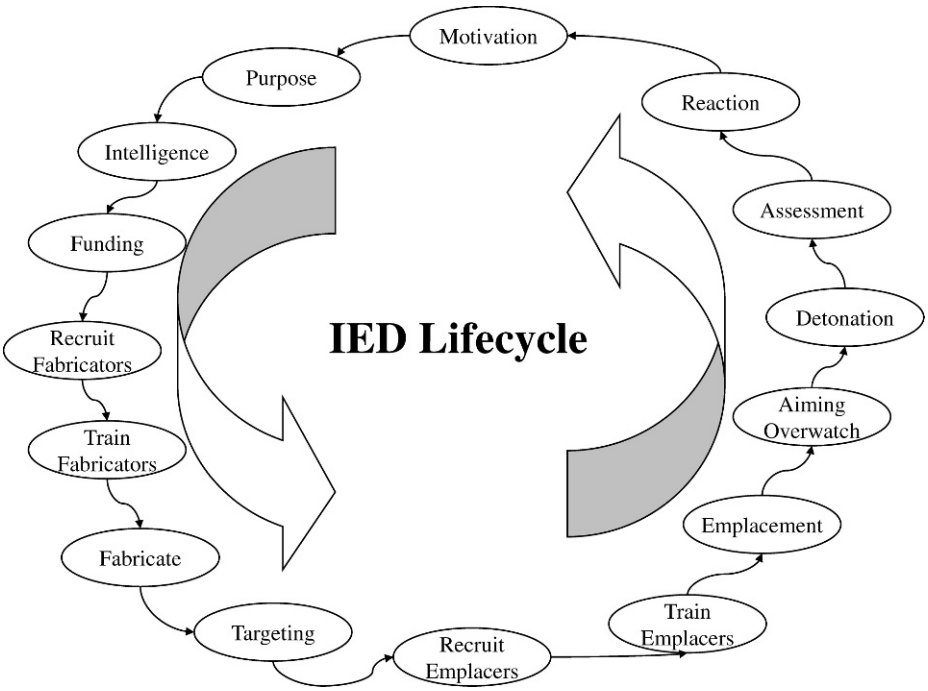

Figure 5. IED lifecycle.

developed NII components *(Figure 3)*. Installation of a variety of physical NII sensors; intermediate data collection and processing units; time, space, positioning information (TSPI) devices, and communications (telemetry) units must be performed on targets (IEDs), and detection platforms (e.g., explosive ordinance disposal [EOD] robots, fixed and maneuverable autonomous surveillance platforms, combat engineers, and EOD vehicles). All newly developed NII components installed are "smart" and can self-identify their capabilities using metadata to the test data acquisition unit. Legacy components must be defined manually and configured using master–slave metadata wrapping concepts. The DAS reconfigures all discovered and configured legacy elements to form a synthetic situational assessment sensor. The synthetic sensor uses data from all component sensors augmented with workflow, algorithmic, spatial, temporal, and contextual metadata to construct the virtual measurement in real time.

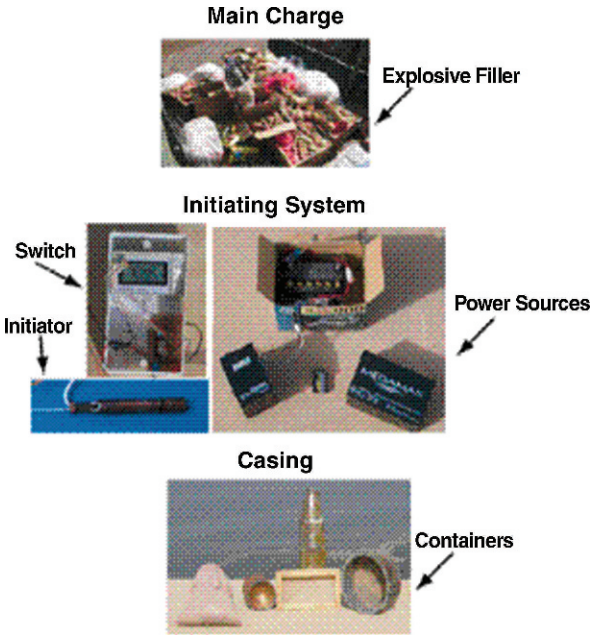

Figure 6. Components of an IED.

The configuration requires the use of sensor and test planning tools and configuration metadata available through vendors' metadata (e.g., TEDs, TransducerML, and SensorML) using a common set of standards for both hardware and software. Once configured and the test commences, sensor measurements are collected and tagged with appropriate metadata (e.g., time tag, location, pedigree, provenance, etc.) indicating all relevant test data for the DAS to use and for storage for future reference and use. It is assumed that none of the sensors is physically interfaced with a SUT devices' data bus, though through observational services, test engineers can have native generated measurement data made available for consumption outside of the SUT operational envelope. The DAS network composed of all relevant source and sink data collection sites is self-configuring. The DAS network provides additional services to ensure self-

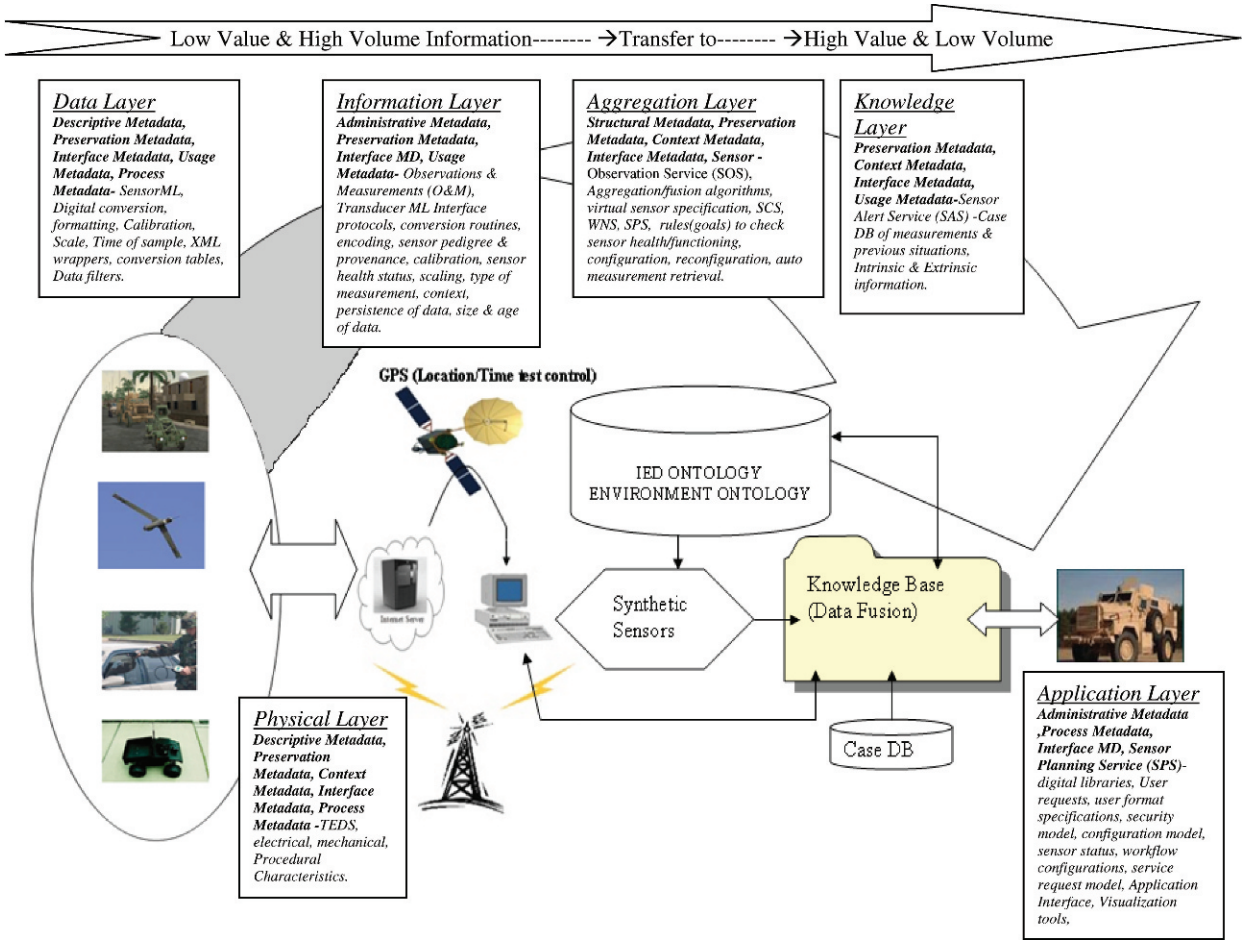

Figure 7. Control flow model and metadata use.

reliability through reconfiguration services and redundancy, which are specified and configured into the devices at initial DAS configuration time (using metadata). All collected data are stored redundantly on board each collection site SUT and in a composite repository off board.

Planning and test preparation phase

During the planning phase, the test engineer uses test planning services that include standardized tools for sensor planning (e.g., sensor web enablement [SWE] sensor planning service (SPS) services) *(Figure 7)*. The tools are used to manage requests by applications to plan events (e.g., set fixed optical sensors to take images at x frames per second, beginning at time t_i and ending at t_{i+j} to synchronize with scheduled IED firing event), configure SUT sensor resources, reconfigure resources (e.g., reset target IED simulators and emulators to provide desired test signatures), calibrate sensors (e.g., test sensitivity, resynchronize time clocks), detect or find sensors (e.g., request inventory of all available sensors, services available, and present status), support addition of new sensors, and initiate the collection and dissemination of measured data.

In the IED detection system test, a test engineer desires to construct a synthetic measurement from available measurements found on the detection mobile vehicle, an EOD talon robot, an AUV, and multiple fixed sensors of varying type and pedigree. The system requires context metadata (e.g., environmental conditions, ontological description of the test domain including GPS signatures from all in situ fixed sensors, derivable through open web services [OWS] and SWE services augmented with Open Grid Forum [OGF] and Open Ontology specification tools), administrative metadata (e.g., security enforcement for application construction and data use, access, and integration services for legacy wrapper access and builds, localization, and authorization for real and virtual sensor configuration historical files for use in configuration builds), descriptive metadata (e.g., specific sensor data used to define real and virtual sensor state such as TEDs, TransducerML, SensorML, as well as work

flow specifications for building and operating synthetic instruments), and process metadata to locate sensors, algorithms, and distribute the tasks, based on priority, needed to build the user defined synthetic instruments from basic measurements and known sensors. Once synthetic sensor specification is completed, available resources must verify their capabilities to accomplish the task using structural metadata (e.g., configuration checks, status checks, and calibration checks) and preservation metadata (e.g., location and configuration of sources, sinks, and paths) and descriptive metadata elements.

Any required transformations will also be defined at this time using interface metadata (e.g., correlating map data with geographic information systems [GIS] data and TSPI data, along with sensor streams to facilitate real-time streaming synthetic sensor operations) and placed in the structural specification for the synthetic sensor. Many system level services and service oriented applications must be developed to allow for the actual control (e.g., error detection, error correction) of the system during the test.

Test execution phase

In the IED defeat system test example, a trace for a unique sensor measurement through two faces of the EISA model, the information and the control models, is performed. Within the examination is a definition for raw sensor measurements flow from a sensor to the monitoring applications, illustrating metadata extracted and used for interpreting measurements from the source through the application sink.

Raw measurements are detected and verified by a sensor (described through descriptive metadata such as SensorML), stored (using preservation metadata, e.g., Structured Query Language, Open Web Language), extracted and translated into information (using interface metadata, TransducerML, SensorML, sensor observation services (SOS)). The extracted information is also marked with usage tags and contextual tags (using extracted usage and context metadata, including provenance and pedigree information to support replay, query, or restoration). These added metadata initiate data provenance and pedigree chain formations for future data preservation. In the IED detection example, the extracted information is combined with additional measurement, spatial, and contextual information items using structural and process metadata to build a synthetic sensor measurement supporting the multisensor detection scenario. To perform data fusion and aggregation operations using multiple heterogeneous data fragments as inputs, developers must extract process metadata describing the fusion algorithm(s) to utilize (possibly using Sensor ML) or the dataflow processes to use a structural metadata item defining a synthetic sensors specification (also possibly using Sensor ML). Aggregated metadata are used along with context metadata to place the appropriate domain specific parameters to aid the goal directed synthetic computation (using TransducerML and Ontological specification). Context metadata, in the form of domain ontology, are used to place the synthetic data in a place and time for the IED defeat system and component under test.

The derived composite synthetic, aggregated measurement is then transferred to the knowledge layer to be used for additional actionable knowledge development. This may include incorporation into the knowledge base as a new case instance or an ontological instance. Transport metadata are used to aid in payload specification, packaging, and transmission. Along the synthetic sensors' data flow path (Figure 8), each metadata item extracted and used in the transformation of the raw measurement into actionable knowledge is tagged and stored using usage metadata, preservation, and administrative metadata supporting provenance storage, long term management, and postretrieval of the measurements.

The stored ontological and instance information is used to maintain lineage and pedigree of measurements. Using preservation metadata, the derived synthetic situational measurement is persistently stored, maintaining information such as how, what, where, why, and by whom was this measurement stored. Preservation information can later be used to retrieve and restore measurements for future uses. Along the dataflow path *(Figure 8)*, administrative metadata are used to determine if the application requesting the measurement has appropriate authorizations and to orchestrate the performance of predefined workflows (process metadata) performing desired actions.

Another important service provided is support for the configuration and logging services to aid in the collection of build information into usage metadata for future pedigree and provenance determination. The example illustrates a small fragment of the services needed to support a user's construction of an IED defeat system test scenario, from the initial definition of all sensor and service resources available, selection of desired resources, to configuration into desired analysis displays.

Standards for interoperability and reuse

No single standard seems to capture all the elements needed within the IED defeat system example test scenario. To capture requirements to support placement of items in a map coordinate system and within a

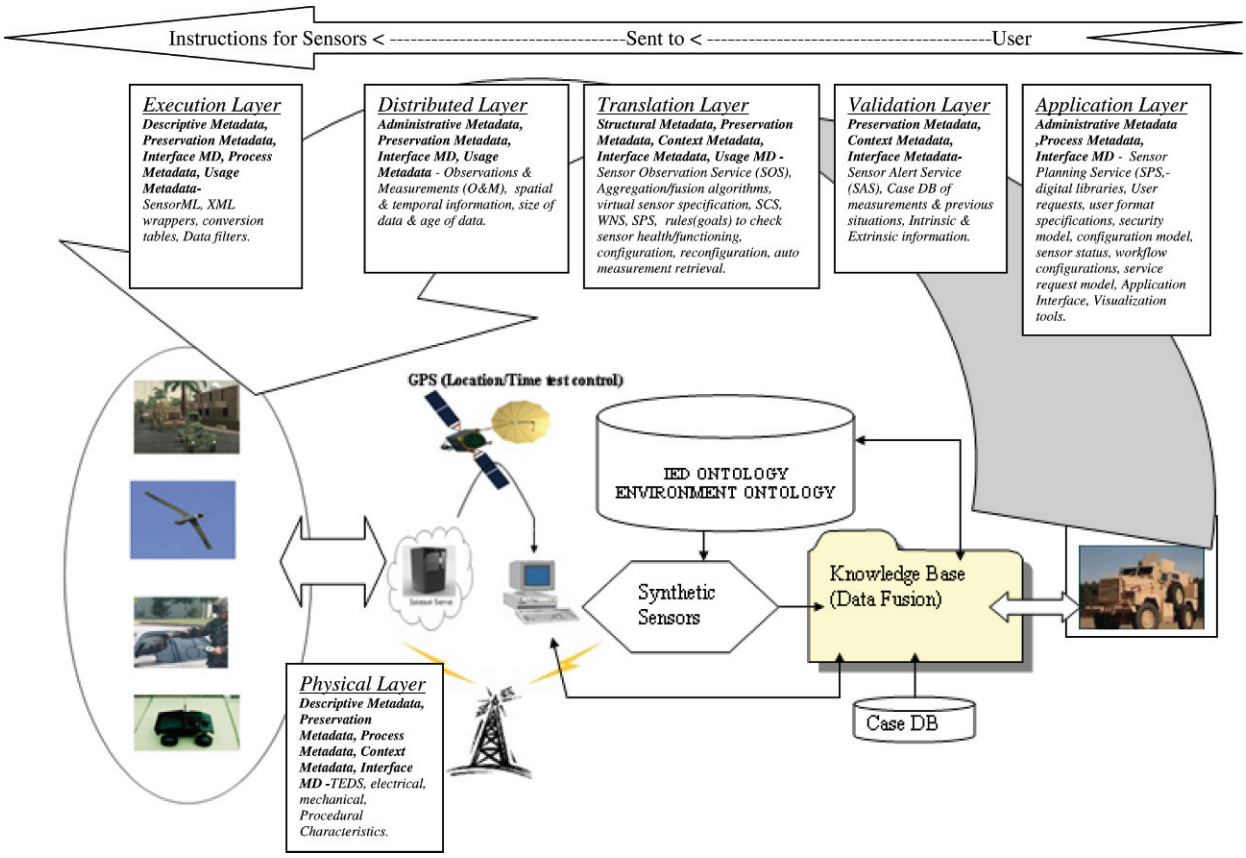

Figure 8. Information flow and metadata use.

space and time context, we require one set of services. Another set of services are needed to use these specifications to locate resources (sensors and services) to configure these services for a specific planned use, to control their operation to effectively extract the correct observations, to configure and control alerts allowing for real-time control of boundaries or events of interest, and to correctly capture data and provenance and pedigree metadata for use in postanalysis or modeling.

Figure 9 depicts a plausible configuration of available standards and services that would be needed to plan, build, configure, operate, and analyze the scenario we have postulated. The Open Geospatial Consortium (OGC) and the open grid forum are collaborating on collections of open standards that address many of the distributed computing and geospatial issues required by the testers in building distributed tests for systems of systems testing, such as is found in the IED detection system described within this article.

The OGC (2004) has specified a set of web services (OGC Web Services or OWS) standards that can be layered on top of sensor specific services to provide distributed geospatial services. One of the standards within this collection is the Web Feature Service, which provides standards for retrieval and update of digital representations of real-world entities tied to the earth's surface. The Web Mapping Service standardizes the integration and display of superimposed map

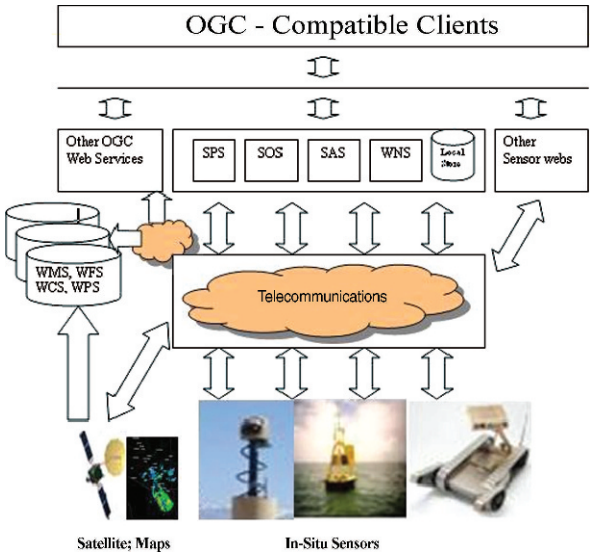

Figure 9. Available and evolving standards for example test use.

entities from multiple heterogeneous sources, the Web Coverage Services standardize access to spatially extended coverages (details), and the Web Processing Services provide standardized basic request and response interaction protocols and metadata formats for remote execution of any algorithm, calculation, or model operating on spatial data. The Catalogue service provides standardized services and metadata specifications to publish, discover, browse, and query metadata about information and services available.

The OGC's SWE standards provide distributed and local services for accessing, controlling and using sensors, instruments and imaging devices. The SWE standards consist of four primary standards: Sensor Planning Service, Sensor Observation Services, Sensor Alert Services, and Web Notification Service. The Sensor Planning Service standardizes the tasking of sensors or models.

Tasks include reprogramming, calibrating, starting, altering a sensor mission, and controlling simulation models. The Sensor Observation Services are used to standardize the methodology and metadata used in retrieval of measurement observations from a sensor or model. Included in the retrieved information is sensor system configuration and status metadata information. Sensor Alert Services provide standard protocols, interfaces, and metadata specifications for subscribing to or publishing alerts configured for sensors. The Web Notification Service defines a set of standard specifications that control and configure the way Web Services interact using a predefined notification collection or pattern.

The low level sensors also require additional standardized services for specifying individual instances of a sensor and models including workflows defining composite sensors and synthetic sensors, spatial locations, contextual information, and other relevant information that will make them available to the upper level services. The basic standards available for such descriptions include, but are not limited to, the IEEE 1451 standards, TEDS, TransducerML, and SensorML standards. These standards are further augmented with ontological standards for OWL (McGuiness 2004) and database standards (e.g., SQL database language [Melton 2003] or OMG—object database or metaobject facility [Pope 1998]) to provide tools to organize these resources into persistent collections of device specifications, measurement instances, and knowledge repositories.

Conclusion

In this article, a definition of metadata and how it is used in the context of sensor systems applied to T&E is developed. The term "metadata" is found to be more than simply "data about data." "Metadata" has many meanings and many applications dependent on the context. The type of metadata is not uniform, but instead is defined based on a few fundamental questions, what is it, how is it used, where is it used, who generated it, how is it related to other data. In this article, we developed these concepts and applied them to an example configuration, control, and execution thread within a heterogeneous distributed multisensor IED defeat system T&E scenario. The natural conclusion from this effort is to look toward metadata and sensor network service standards to realize the true value and strength of metadata to enhance seamless communication and interoperability and reuse of sensor and NII collected information for the T&E community. ❑

Paul J. Fortier, Ph.D., is a professor of electrical and computer engineering at the University of Massachusetts Dartmouth located in North Dartmouth, Massachusetts. He is a member of the IEEE and ACM as well as numerous technical and working groups of these societies. He is a past member of the SQL Standards Committee, the Predictable Real-time Systems Task Group, the ANSI POSIX Working Group, and present member of the American Medical Informatics Associations HL7 EHR Standards Committee and the iNet Metadata, Test Article and Systems Management Standards Working Groups. Prior to his arrival at UMass Dartmouth, he served as a senior systems engineer at the Naval Undersea Warfare Center for 16 years. His research interest are in real-time and temporal databases, data mining and knowledge discovery, medical informatics, real-time and semantic concurrency control and transaction processing models, performance modeling, database language standards, and embedded sensor systems and sensor networks. E-mail: pfortier@umassd.edu

Kiran Dasari is a graduate student in the Electrical and Computer Engineering Department at the University of Massachusetts at Dartmouth. He will complete requirements for his master of science in computer engineering in December of 2008. Mr. Dasari's interests are in business process models and in metadata models for ontological data representation. He will be joining the Ross School of Business at the University of Michigan, Ann Arbor in January of 2009. He has plans to set up his own consulting firm in emerging economies upon completion of his advanced degree.

References

Baca and Murtha, ed. 1998. *Introduction to metadata: Pathways to digital information*. Washington, DC: Getty Information Institute. Version 2.1. http://

www.getty.edu/research/conducting_research/standards/intrometadata/index.html (accessed June 10, 2008).

Biagioni, E. and K. Bridges. 2002. The application of remote sensor technology to assist the recovery of rare endangered species. *International Journal of High Performance Computing Applications*. 16 (3): 315–324.

Borst, W. N. 1997. Construction of engineering ontologies, Doctoral thesis, University of Twente, Enschede.

Bose, R. 2002. A conceptual framework for composing and managing scientific data lineage. In *14th international conference on scientific and statistical database management*, ed. J. Kennedy. 15–19. Edinburgh, Scotland: IEEE Computer Society.

Botts, M., ed. November 2004. *Sensor Model Language (SensorML) for in-situ and remote sensors*. Location: Publisher. http://vast.nsstc.uah.edu/SensorML/Sensor ML_04-019_1.0_beta.pdf (accessed October 20, 2007)

Committee on Catalog Description and Access (CCDA). 2000. *Task force on metadata: final report* (CC:DA/TF/Metadata/5), June 16. http://www.libraries.psu.edu/tas/jca/ccda/tf-meta6.html (accessed June 16, 2008).

Dasari, K. 2008. Development of sensor ontology for environmental monitoring and situational assessment. Master's thesis, University of Massachusetts Dartmouth.

Duval, Erik, Wayne Hodgins, Stuart Sutton, and Stuart L. Weibel. 2002. Metadata principles and practicalities. *D-Lib Magazine*. 8 (4): 16 pp.

Faulstich, R. and T. Grace. 2007. Integrated Network Enhanced Telemetry (iNet) overview. In *Proceeding of the 38th Annual International Symposium of the Society of Flight Engineers,* October 22–25, 2007/ Las Vegas, NV. Washington, DC: ITC.

Fortier, P. and K. Dasari. 2008. The role of metadata for improving the utility and life of sensor derived measurement. In *ITEA Annual Technology Conference,* July, Colorado Springs, CO, 10 pp. Washington DC: ITEA.

Gauvin, M., A. Boury-Brisset. and A. Auger. 2004. Context, ontology and portfolio: key concepts for a situational awareness knowledge portal. In *Proceedings of the 37th Hawaii International Conference on Systems Sciences,* 2004 January 5–8, 2004, Hawaii, HI, 10 pp. New York: IEEE.

Groth, P., Michael Luck, and Luc Moreau. 2004. A protocol for recording provenance in service-oriented grids. In *Proceedings of the 8th International Conference on Principles of Distributed Systems* (OPODIS'04), December 2004, Grenoble, France, 124–139. New York: Springer Verlag.

Gruber. and Thomas, R. 1993. A translation approach to portable ontology specifications. *Knowledge Acquisition*. 5 (2): 199–220.

Hall, D. and J. Llinas. 1997. An introduction to multisensor data fusion. *Proceedings of the IEEE*. 85 (1): 6–23.

Harpel, D. 2007. Shadow 200 TUAV proves mettle in heavy Iraq use. *Defense Systems Daily*. November 6.

Havens, S. (ED), 2007. *Transducer Markup Language Implementation Specification*. Wayland, MA: Open GIS Consortium. Available at http://www.opengeospatial.org/ standards (accessed April 15, 2009).

Joshi, H. and H. Michel. 2007. Integrating information-centric, control-centric and behavior-centric technical reference models for autonomous sensor networks *ICWN*. 2007: 319–324.

Kauchak, M. July 2006. New eye in the sky *Journal of Military Geospatial Technology*. 4 (3): 5 pp.

Ladner, R. and Frederick E. Pe. 2005. Metadata concepts to support a net-centric data environment. In *Net-centric approaches to intelligence and national security*. 29–54. New York: Springer Verlag.

Ledlie, J., Chaki, Ng, David A. Holland, Kiran-Kumar Muniswamy-Reddy, Uri Braun, and Margo Seltzer. 2005. Provenance-aware sensor data storage. In *Proceedings of First IEEE International Workshop on Networking Meets Databases*. April 8–9, 2005, Tokyo, Japan, 10 pp, New York: IEEE.

Lee, C. and Percivall, G. 2008, November. Standards based computing capabilities for distributed geospatial applications *IEEE Computer*. 41 (11): 50–66.

Lee, K. 2006a. Sensor networks pilots for DRM 2.0: Sensor standards harmonization working group meeting, September 12, 2006. http://colab.cim3.net/file/work/SIC0p/2006-04-12 (accessed April 14, 2009).

Lee, K. 2006b. An implementation of the proposed IEEE 1451.0 and 1451.5 standards. In *Proceedings of the Sensors Applications Symposium,* February 5–7, 2006, Houston, TX. 72–77. New York: IEEE.

MacIntyre, A. 1972. Ontology. In *The encyclopedia of philosophy*, ed. Paul Edwards. volume 5, 542–543. New York: Macmillan Publishing and The Free Press.

Mainwaring, A., et al. 2002. Wireless sensor networks for habitat monitoring. In *ACM International Workshop on Wireless Sensor Networks and Applications,* September 2002. September 28, 2002. Atlanta, GA, 8 pp. New York: ACM.

McGuiness and Van Harmelen (editors). 2004. Wayland, MA: Open GIS Consortium. http://www.omg.org/ standards (accessed October 20, 2008).

Melton, J. (editor). 2003. SQL Framework. New York: IEEE Society. http://www.jcc.com/sql.htm (accessed June 14, 2008).

Michel, H. and P. Fortier. May 2006. Development of an embedded instrumentation system architecture and its comparison to the test and training enabling architecture. In *Proceedings of the SPIE Symposium on Defense and Security*. December. Las Vegas, NV: SPIE.org. 7p.

NISO. 2004. *Understanding Metadata*. Bethesda, MD: NISO Press. http://www.niso.org/standards/resources/UnderstandingMetadata.pdf (accessed June 10, 2008).

Nle, N., U. Ppeler, D. Nicklas, T. Schwarz, and M. Grossmann. 2005. Benefits of integrating meta data into a context model. In *Proceedings of the 3rd International Conference on Pervasive Computing and Communications Workshop*. March 8–11, 2005, Kavai, HI. 12 pp. New York: IEEE Publishing.

Noseworthy, J. R. 2005. Developing distributed applications rapidly and reliably using the TENA middleware. In *Proceedings of the IEEE Military Communications Conference (MILCOM),* 7p. October 17–20, 2005, Atlantic City, NJ. 7pp. New York: IEEE Publishing.

OGC (Open Geospatial Consortium) 2008. Open GIS Specifications. Wayland, MA: Open GIS Consortium. www.opengeospatial.org (accessed June 5, 2008).

Park, S., J. Kim, K. Lee, K. Shin, and D. Kim. 2006. Embedded sensor networked operating system. In *Ninth IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing*. April 24–26, 2006, Gyeongju, South Korea, pp. 117–124. Washington DC: IEEE.

Pope. 1998. The COBRA Reference Guide. Waltham, MA: Addison Wesley Publisher. http://www.w3.org/2004/OWL/ (accessed June 14, 2008).

Qin, J. and J. C. Prado. 2006. The semantic and syntactic model of metadata. In *Alfabetização Digital e Accesso ao Conhecimento (Série Comunicação da Informacao Digital 2006)*. 4: 143–156.

Sethunadh, R., S. Athuladevi, and S. Iyer. October. 2002. Virtual instrumentation techniques in test and evaluation of launch vehicle *Avionics, Defense Science Journal*. 52 (4): 357–362.

Tambouris, E., N. Manouselis. and C. Costopoulou. 2007. Metadata for digital collections of e-government resources *The Electronic Library (TEL), Special Issue on Metadata and Semantics for Digital Libraries and Information Centres*. 25 (2): 176–192.

Tan, V. H. K. 2004. *Interaction tracing for mobile agent security*. Doctoral thesis, University of Southampton: England.

Tannenbaum, A. 1998. The metadata mystique. *The Journal of Data Warehousing*. 3 (4): winter.

Visnevski, N. 2008. Embedded instrumentation systems architecture. In *Proceedings of the IEEE Instrumentation and Measurement Technology Conference. IMTC 2008.* May 12–15, 2008, Lancaster, CA. pp. 1134–1139. Washington, DC: IEEE Publishing.

Visnevski, N. and T. Johnson. 2007. Embedded instrumentation systems architecture for mobile wireless sensing platforms. *ICWN*. 2007: 335–342.

Wang, H., J. Elson, L. Girod, D. Estrin, and K. Yao. 2003. Target classification and localization in habitat monitoring. In *Proceedings of the IEEE ICASSP,* April 6–8, 2003, Los Angeles, CA. pp. 844–847. Washington, DC: IEEE Publishing.

Acknowledgment

The authors thank the Test Resource Management Center (TRMC) Test and Evaluation/Science and Technology (T&E/S&T) program for their support. This work was partially supported by the T&E/S&T Program through the Naval Undersea Warfare Center, Newport, Rhode Island, under NUWCDIVNPT SEAPORTE contract N00178-04-D-4109-N407. We also wish to thank Purvis Systems Incorporated of Middletown, Rhode Island, for their support under ATMC work order number 080825-1 through the Advanced Technology and Manufacturing Center of the University of Massachusetts Dartmouth.